International Journal on Advanced Science Engineering Information Technology

Vol. 2 (2012) No. 5
ISSN: 2088-5334

Computational Analysis of PTEN Gene Mutation

Siew-Kien Mah [#], Sim-Hui Tee [*]

[#] *Faculty of Engineering, Multimedia University, 63100 Cyberjaya, Malaysia*
E-mail:skmah@mmu.edu.my

**Faculty of Creative Multimedia, Multimedia University, 63100 Cyberjaya, Malaysia*
E-mail: shtee@mmu.edu.my

Abstract— **Post-genomic data can be efficiently analyzed using computational tools. It has the advantage over the biochemical and biophysical methods in term of higher coverage. In this research, we adopted a computational analysis on PTEN gene mutation. Mutation in PTEN is responsible for many human diseases. The results of this research provide insights into the protein domains of PTEN and the distribution of mutation.**

Keywords— **Gene mutation, PTEN, genomic data.**

I. Introduction

Computational tools are efficient approaches in the analysis of large volume of post-genomic data. There exists a wide range of computational tools [1-6] for genomic analysis and protein analysis, all of which built upon powerful algorithms. Without exception, there is an increasing trend of applying computational analysis for gene discovery and treatment. In this research, we adopted a computational approach in the analysis of Phosphatase and Tensin homolog (PTEN) gene mutation.

PTEN is a tumor suppressor gene which has been implicated in many diseases, including cancer. It involves in a wide range of physiological processes by negatively regulating the PI3K-Akt signalling pathway [7,15]. Activation of Akt results in the instability of PTEN and consequently induces drug resistance to cetuximab and gefitinib [11]. Recent study has shown that PI3K-PTEN signalling cascade is important in protecting cells against oxidative stress [8]. Besides, it was observed that PTEN might reverse chemoresistance to cisplatin and may be targeted for molecular treatment of ovarian cancer [9]. Clinical studies have found that an increase in PTEN expression level is correlated to longer survival [12] in certain cancers, such as extrahepatic cholangiocarcinoma [10]. Although a diverse implication of PTEN has been discovered in cancer therapy, there are more researches needed to carry out to study the impact of PTEN in molecular treatment.

Mutation in PTEN is implicated in many human diseases [13]. A review done by Tainsky demonstrates that germline mutation in PTEN causes more than 10 types of cancer in human [16]. The understanding of individual genetic mutation is important for a better prediction in disease treatment [14]. In this study, we attempted to analyze PTEN gene mutation using computational approach. Analysis on RNA transcripts and protein domains are included. The results of this study would provide *in silico* insights to the mutation in PTEN.

II. Methods

We used COSMIC database [2] to mine the somatic mutation information of PTEN gene. COSMIC is a public database which curates information on somatic mutations in cancer and links to external data sources such as Ensembl and The Cancer Genome Atlas Project (TCGA) [2]. We used the latest version of COSMIC, which is version 53 as at May 2011, with the last update in March 2011. We identified the mutations in term of substitution, insertion and deletion. Both cDNA and amino acid sequence type of mutation distribution are identified.

PTEN signalling pathway was investigated. In addition, we identified and analyzed protein domain for PTEN using Pfam [17] and InterPro [18]. We performed clustal alignment for the protein domains. Lastly, we used Cn3D version 4.3 to model the protein tertiary structure.

III. Results and Discussion

The COSMIC shows that there are 16169 unique samples of PTEN gene in human genome, of which only 2005 are mutated samples. The histology of cancer implicated by

PTEN includes carcinoma, glioma, hyperplasia, neoplasm, sarcoma and melanoma. Some of the PTEN-implicated cancer samples are provided in Table 1.

TABLE 1
A partial list of PTEN-implicated cancer samples

COSMIC sample ID	Amino acid	Primary tissue	Histology	Mutation ID
1009627	p.R130*	ovary	Carcinoma	21342
1009628	p.R233*	endometrium	Carcinoma	21343
1010524	p.L57S	CNS	Glioma	5127
1041979	p.G230E	endometrium	hyperplasia	23550
1047275	p.D116G	thyroid	Carcinoma	23662
1229477	p.I253N	skin	Melanoma	5230
848117	p.T401I	soft tissue	Leiomyo-sarcoma	5124

The diverse histology and primary tissue involved in PTEN gene mutation implies that PTEN is expressed in multiple organs and being up-regulated/down-regulated in multiple signalling pathways. The mutation overview chart of PTEN is shown in Fig. 1.

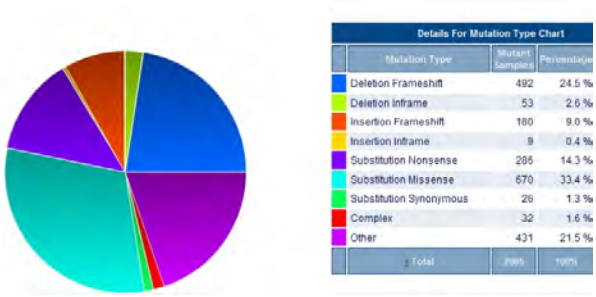

Fig. 1 Mutation overview chart of PTEN

Fig. 1 shows that gene substitution constitutes the main mutation of PTEN, where it weighs 49% of total mutation. Gene deletion consists of 27.1%. The total percentage of indels is only 36.5%. Fig. 2 shows the chart for gene substitution.

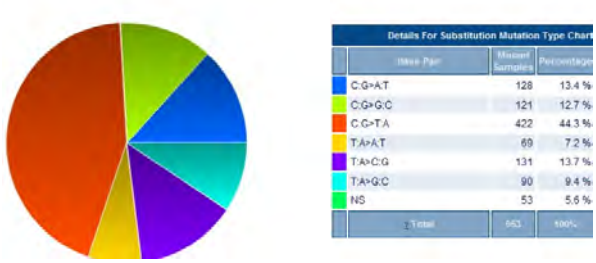

Fig. 2 Substitution chart

There are two types of substitution. The most common type for point mutation is the transition, in which the substitution takes place between the same types of nitrogenous base; another type of substitution is the transversion in which the substitution takes place between different types of nitrogenous base. Fig. 2 shows that 44.3% of substitution is CG pair to TA pair, which is a transition substitution that comprises the largest group. TA pair to CG pair substitution, which constitutes 13.7% of overall substitution is also a transition. Transversion substitution consists of 42.7% of overall substitution; whereas nonsense substitution consists of 5.6%. The distribution of somatic mutation is given in Fig. 3. The figure shows that substitution is highly concentrated at amino acid sequence 120-130, 160-170, and 220-230; whereas indels occur mainly at the sequence <350.

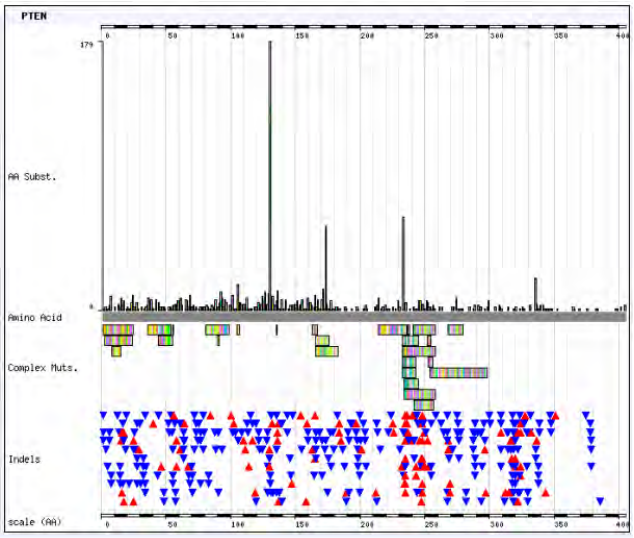

Fig. 3 The distribution of somatic mutation of PTEN

Fig. 4 shows that PTEN is implicated in T-cell receptor (TCR) signalling pathway, one of the signalling pathways in human immune system. The figure shows that PTEN is implicated in the phosphorylation of other substrates. It implies that PTEN is important in the control of the activity of various enzymes [19, 21], which is one of the features of phosphorylation, in the immune system. In addition, it was reported that Akt could be activated through the phosphorylation of erythropoietin receptor [20]. This finding sheds lights for geneticists to study in detail the connection between PTEN mutation and phosphorylation.

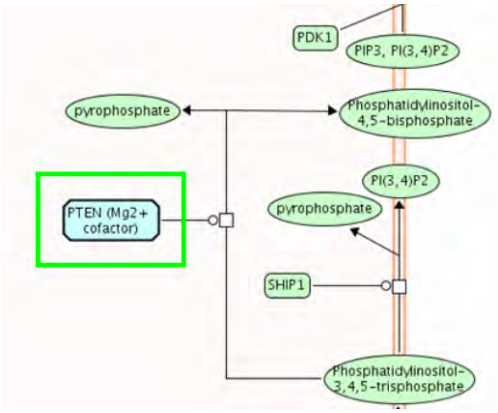

Fig. 4 The implication of PTEN in TCR signalling pathway

We identified and analyzed protein domains for PTEN using Pfam [17] and InterPro [18]. Using Pfam, we obtained two domain families for PTEN (P60484), which are DSPc and PTEN_C2. DSPc starts from position 47 and ends at position 175; whereas PTEN_C2 starts from position 188 and ends at position 349. The total length of PTEN protein is 403 amino acids. InterPro was used to analyze PTEN_C2 domain. It showed that this domain matches 432 proteins in human proteome. This domain functions in protein binding, as shown by the Gene Ontology annotation.

The human PTEN domain was aligned with mouse model, as shown in Fig. 5. The highlighted blocks in yellow represent the aligned domain of PTEN between these two organisms. It shows that PTEN is largely conserved in both organisms.

```
1    MTAIIKEIVSRNKRRYQEDGFDLDLTYIYPNIIAMGFPAERLEGVYRNNIDDVVRFLDSK  60 P60484 PTEN_HUMAN
1    MTAIIKEIVSRNKRRYQEDGFDLDLTYIYPNIIAMGFPAERLEGVYRNNIDDVVRFLDSK  60 O08586 PTEN_MOUSE
     ************************************************************

61   HKNHYKIYNLCAERHYDTAKFNCRVAQYPFEDHNPPQLELIKPFCEDLDQWLSEDDNHVA 120 P60484 PTEN_HUMAN
61   HKNHYKIYNLCAERHYDTAKFNCRVAQYPFEDHNPPQLELIKPFCEDLDQWLSEDDNHVA 120 O08586 PTEN_MOUSE
     ************************************************************

121  AIHCKAGKGRTGVMICAYLLHRGKFLKAQEALDFYGEVRTRDKKGVTIPSQRRYVYYYSY 180 P60484 PTEN_HUMAN
121  AIHCKAGKGRTGVMICAYLLHRGKFLKAQEALDFYGEVRTRDKKGVTIPSQRRYVYYYSY 180 O08586 PTEN_MOUSE
     ************************************************************

181  LLKNHLDYRPVALLFHKMMFETIPMFSGGTCNPQFVVCQLKVKIYSSNSGPTRREDKFMY 240 P60484 PTEN_HUMAN
181  LLKNHLDYRPVALLFHKMMFETIPMFSGGTCNPQFVVCQLKVKIYSSNSGPTRREDKFMY 240 O08586 PTEN_MOUSE
     ************************************************************

241  FEFPQPLPVCGDIKVEFFHKQNKMLKKDKMFHFWVNTFFIPGPEETSEKVENGSLCDQEI 300 P60484 PTEN_HUMAN
241  FEFPQPLPVCGDIKVEFFHKQNKMLKKDKMFHFWVNTFFIPGPEETSEKVENGSLCDQEI 300 O08586 PTEN_MOUSE
     ************************************************************

301  DSICSIERADNDKEYLVLTLTKNDLDKANKDKANRYFSPNFKVKLYFTKTVEEPSNPEAS 360 P60484 PTEN_HUMAN
301  DSICSIERADNDKEYLVLTLTKNDLDKANKDKANRYFSPNFKVKLYFTKTVEEPSNPEAS 360 O08586 PTEN_MOUSE
     ************************************************************

361  SSTSVTPDVSDNEPDHYRYSDTTDSDPENEPFDEDQHTQITKV  403  P60484   PTEN_HUMAN
361  SSTSVTPDVSDNEPDHYRYSDTTDSDPENEPFDEDQHSQITKV  403  O08586   PTEN_MOUSE
     ************************************:*****
```

Fig. 5 PTEN domain alignment between human and mouse model

We then performed an alignment between human and mouse PTEN for mutagenesis. The result is shown in Fig. 6, which clearly demonstrates that PTEN gene mutates differently in human and mouse. This implies that the diseases caused by PTEN mutation are unlikely to be the same in human and mouse. The aligned point mutation is highlighted in blue.

```
1    MTAIIKEIVSRNKRRYQEDGFDLDLTYIYPNIIAMGFPAERLEGVYRNNIDDVVRFLDSK  60 P60484 PTEN_HUMAN
1    MTAIIKEIVSRNKRRYQEDGFDLDLTYIYPNIIAMGFPAERLEGVYRNNIDDVVRFLDSK  60 O08586 PTEN_MOUSE
     ************************************************************

61   HKNHYKIYNLCAERHYDTAKFNCRVAQYPFEDHNPPQLELIKPFCEDLDQWLSEDDNHVA 120 P60484 PTEN_HUMAN
61   HKNHYKIYNLCAERHYDTAKFNCRVAQYPFEDHNPPQLELIKPFCEDLDQWLSEDDNHVA 120 O08586 PTEN_MOUSE
     ************************************************************

121  AIHCKAGKGRTGVMICAYLLHRGKFLKAQEALDFYGEVRTRDKKGVTIPSQRRYVYYYSY 180 P60484 PTEN_HUMAN
121  AIHCKAGKGRTGVMICAYLLHRGKFLKAQEALDFYGEVRTRDKKGVTIPSQRRYVYYYSY 180 O08586 PTEN_MOUSE
     ************************************************************

181  LLKNHLDYRPVALLFHKMMFETIPMFSGGTCNPQFVVCQLKVKIYSSNSGPTRREDKFMY 240 P60484 PTEN_HUMAN
181  LLKNHLDYRPVALLFHKMMFETIPMFSGGTCNPQFVVCQLKVKIYSSNSGPTRREDKFMY 240 O08586 PTEN_MOUSE
     ************************************************************

241  FEFPQPLPVCGDIKVEFFHKQNKMLKKDKMFHFWVNTFFIPGPEETSEKVENGSLCDQEI 300 P60484 PTEN_HUMAN
241  FEFPQPLPVCGDIKVEFFHKQNKMLKKDKMFHFWVNTFFIPGPEETSEKVENGSLCDQEI 300 O08586 PTEN_MOUSE
     ************************************************************

301  DSICSIERADNDKEYLVLTLTKNDLDKANKDKANRYFSPNFKVKLYFTKTVEEPSNPEAS 360 P60484 PTEN_HUMAN
301  DSICSIERADNDKEYLVLTLTKNDLDKANKDKANRYFSPNFKVKLYFTKTVEEPSNPEAS 360 O08586 PTEN_MOUSE
     ************************************************************

361  SSTSVTPDVSDNEPDHYRYSDTTDSDPENEPFDEDQHTQITKV  403  P60484   PTEN_HUMAN
361  SSTSVTPDVSDNEPDHYRYSDTTDSDPENEPFDEDQHSQITKV  403  O08586   PTEN_MOUSE
     ************************************:*****
```

Fig. 6 Alignment for mutagenesis for human and mouse

The tertiary structure of DSPc domain family of PTEN was modelled using ball and stick representation. Fig. 7 depicts an aligned model where the color represents the aligned pairs; Fig. 8 depicts the domain of DSPc; Fig. 9 depicts the residues of DSPc. A wide range of color implies that DSPc has more than 5 residues; Fig. 10 depicts the identity of sequence conservation; and lastly, Fig. 11 shows that the DSPc domain of PTEN consists of a structure of 6 α -helices and 4 β -sheets. It shows a conformation of helices-loops-sheets, where the loops may serve as a catalytic site in the cellular pathways.

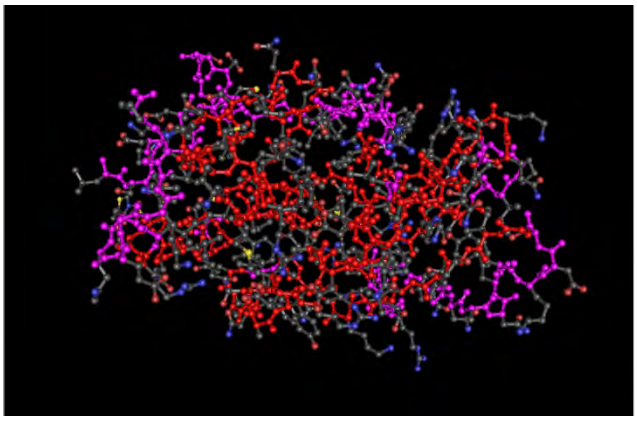

Fig. 7 An aligned model of DSPc

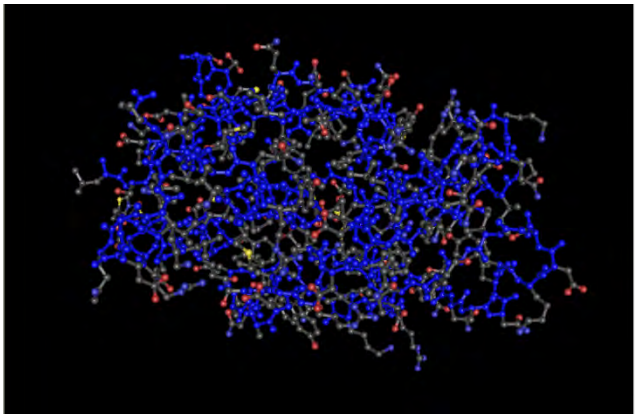

Fig. 8 The domain model of DSPc

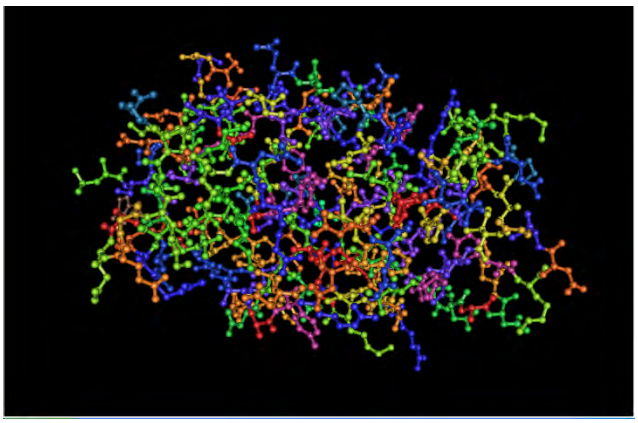

Fig. 9 The residue model of DSPc

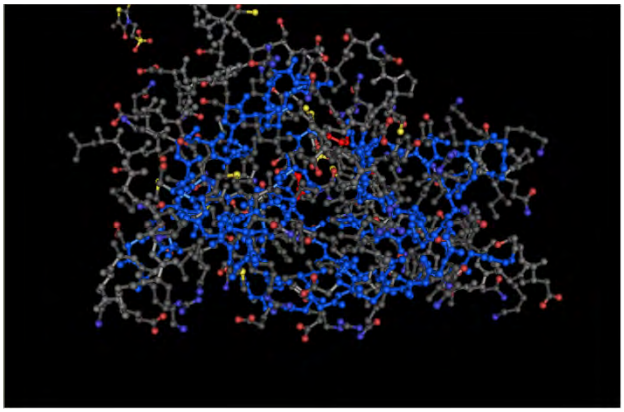

Fig. 10 The identical sequence conservation of DSPc

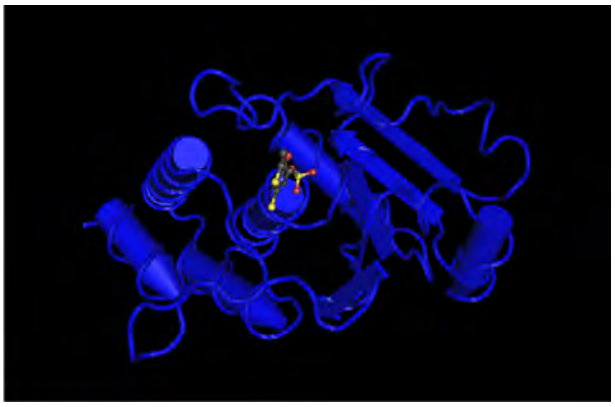

Fig. 11. DSPc domain family consists of 6 α helices 4 β sheets.

IV. Conclusions

Computational analysis of the PTEN gene mutation which is implicated in many human diseases shows that gene substitution constitutes the main mutation (49%) of PTEN. Based on the change in the nucleotide type, the substitution mutation may be classified into transition (CG> TA, TA> CG) and transversion mutations (CG> AT, CG> GC, TA> AT, TA> GC). PTEN regulates various enzymes in the immune system. The human PTEN domain aligned with mouse model is largely conserved but shows differences when alignment is performed for mutagenesis. Molecular model of DSPC domain family of PTEN was also modelled using ball and stick method.

References

[1] J. Ahmed., T. Meinel., M. Dunkel., M. S. Murgueitio., R. Adams., C. Blasse., A. Eckert., S. Preissner., and R. Preissner, "CancerResource: a comprehensive database of cancer-relevant proteins and compound interactions supported by experimental knowledge", *Nucleic Acids Research*, Vol 39, pp D960-D967, 2011.

[2] S. A. Forbes., N. Bindal., S. Bamford., C. Cole., C. Y. Kok, D. Beare., M. Jia., R. Shepherd., K. Leung., A. Menzies., J. W. Teague., P. J. Campbell., M. R. Stratton., and P. A. Futreal, "COSMIC: mining complete cancer genomes in the catalogue of somatic mutations in cancer", *Nucleic Acids Research*, Vol 39, pp D945-D950, 2011.

[3] J. Ponomarenko., H-H. Bui., W. Li., N. Fusseder., P. E. Bourne., A. Sette., and B. Peters, "ElliPro: a new structure-based tool for the prediction of antibody epitopes", *BMC Bioinformatics*, 9: 514, 2008.

[4] A. Martin., M. Ochagavia., L. Rabasa., J. Miranda., J. Femandez-de-Cossio., and R. Bringas, "BisoGenet: a new tool for gene network building, visualization and analysis", *BMC Bioinformatics*, 11: 91, 2010.

[5] W. Stacklies., C. Seifert., and F. Graeter, "Implementation of force distribution analysis for molecular dynamics simulations", *BMC Bioinformatics*, 12: 101, 2011.

[6] M. Kirchner., B. Y. Renard., U. Köthe., D. J. Pappin., F. A. Hamprecht., H. Steen., and J. A. Steen, "Computational protein profile similarity screening for quantitative mass spectrometry experiments", *Bioinformatics*, Vol 26, pp 77-83, 2010.

[7] R. Hou., J. Zhang., T. Yin., H. Cao., N. Zhang., X. Li., L. Wang., Y. Xing., D. Li., and Q. Ji, "Upregulation of PTEN by peroxynitrite contributes to cytokine-induced apoptosis in pancreatic β -cells", *Apoptosis*, Vol 15, pp 877-886, 2010.

[8] K. H. Kang., G. Lemke., and J. W. Kim, "The PI3K-PTEN tug-of-war, oxidative stress and retinal degeneration", *Trends in Molecular Medicine*, Vol 15, pp191-198, 2009.

[9] H. Wu., Y. Cao., D. Weng., H. Xing., X. Song., J. Zhou., G. Xu., Y. Lu., S. Wang., D. Ma, "Effect of tumor suppressor gene PTEN on the resistance to cisplatin in human ovarian cancer cell lines and related mechanisms", *Cancer Letters*, Vol 271, pp260-271, 2008.

[10] J-Y. Chung., S-M. Hong., B. Y. Choi., H. Cho., E. Yu., and S.M. Hewitt, "The expression of phospho-AKT, phosphor-mTOR, and PTEN in extrahepatic cholangiocarcinoma", *Clinical Cancer Research*, Vol 15, pp 660-667, 2009.

[11] S. M. Kim., J. S. Kim., J- H. Kim., C-O. Yun., E. M. Kim., H. K. Kim., F. Solca., S-Y. Choi., and B. C. Cho, "Acquired resistance to cetuximab is mediated by increased PTEN instability and leads cross-resistance to gefitinib in HCC827 NSCLC cells.", *Cancer Letters*, vol 296, pp 150-159, 2010.

[12] F. Solari., A. Bourbon-Piffaut., I. Masse., B. Payrastre., A. M-L. Chan., and M. Billaud, "The human tumour suppressor PTEN regulates longevity and dauer formation in *Caenorhabditis elegans*", *Oncogene*, Vol 24, pp 20-27, 2005.

[13] D. D. Vizio., L. Cito., A. Boccia., P. Chieffi., L. Insabato., G. Pettinato., M. L. Motti., F. Schepis., W. D'Amico., F. Fabiani., B. Tavernise., S. Venuta., A. Fusco., and G. Viglietto, "Loss of the tumor suppressor gene PTEN marks the transition from intratubular germ cell neoplasias (ITGCN) to invasive germ cell tumors", *Oncogene*, Vol 24, pp 1882-1894, 2005.

[14] N. A. Zaghloul., and N. Katsanis, "Functional modules, mutational load and human genetic disease", *Trends in Genetics*, Vol 26, pp 168-176, 2010.

[15] R. Laubenbacher., V. Hower., A. Jarrah., S. V. Torti., V. Shulaev., P. Mendes., F. M. Torti., and S. Akman, "A systems biology view of cancer", *Biochimica et Biophysica Acta: Reviews on Cancer*. Vol 1796, pp 129-139, 2009.

[16] M. A. Tainsky, "Genomic and proteomic biomarkers for cancer: a multitude of opportunities", *Biochimica et Biophysica Acta: Reviews on Cancer*. Vol 1796, pp 176-193, 2009.

[17] R.D. Finn., J. Tate., I. Mistry., P.C. Coggill., S. J. Sammut., H. R. Hotz., G. Ceric., K. Forslund., S. R. Eddy., E. L. sonnhammmer., et. al. "The Pfam protein families database", *Nucleic Acids Research*, Vol 36, pp D281-D288, 2008.

[18] S. Hunter., R. Apweiler., T. K. Attwood., A. Bairoch., A. Bateman., D. Binns., P. Bork., U. Das., L. Daugherty., L. Duquenne., R. D. Finn., J. Gough., D. Haft et. al., "InterPro: the integrative protein signature database", Nucleic Acids Research, Vol 37, pp D211-D215, 2009.

[19] E. L. Huttlin., M. P. Jedrychowski., J. E. Elias., T. Goswami., R. Rad., S. A. Beausoleil., J. Villén., W. Haas., M. E. Sowa., and S. P. Gygi, "A tissue-specific atlas of mouse protein phosphorylation and expression", *Cell*, Vol 143, pp 1174-1189, 2010.

[20] J. Kamishimoto., K. Tago., T. Kasahara., M. Funakoshi-Tago, "Akt activation through the phosphorylation of erythropoietin receptor at tyrosine 479 is required for myeloproliferative disorder-associated JAK2 V617F mutant-induced cellular transformation", *Cellular Signalling*, Vol 23, pp 849-856, 2011.

[21] P. Gruszczyński., M. Obuchowski., and R. Kaźmierkiewicz, "Phosphorylation and ATP-binding induced conformational changes in the PrkC, Ser/Thr kinase from *B. subtilis*", *Journal of Computer-Aided Molecular Design*, Vol 24, pp 733-747, 2010.